End-to-end lossless compression of high precision depth maps guided by pseudo-residual

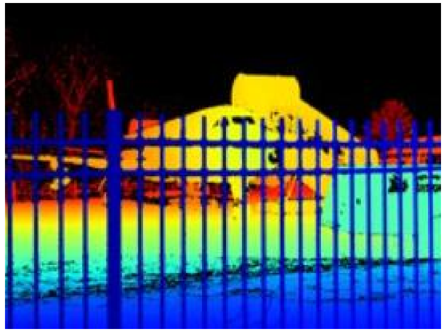
Yuyang Wu, Wei Gao*
DCC 2022

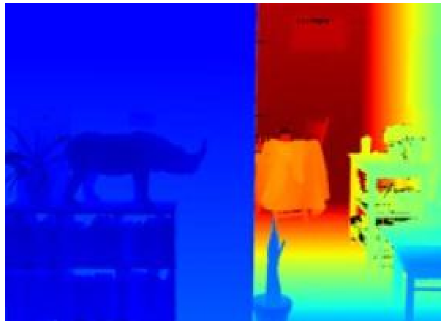
Background

Massive amount of high precision depth maps are produced with the rapid development of equipment like laser scanner or LiDAR.

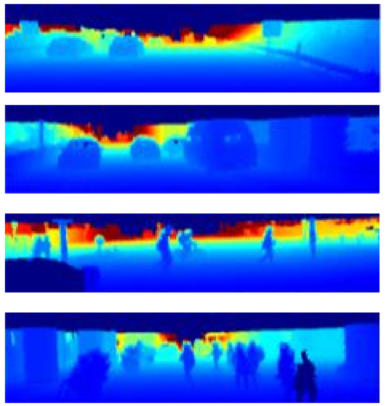
Examples

laser scanner : FARO Focus S350 scanner



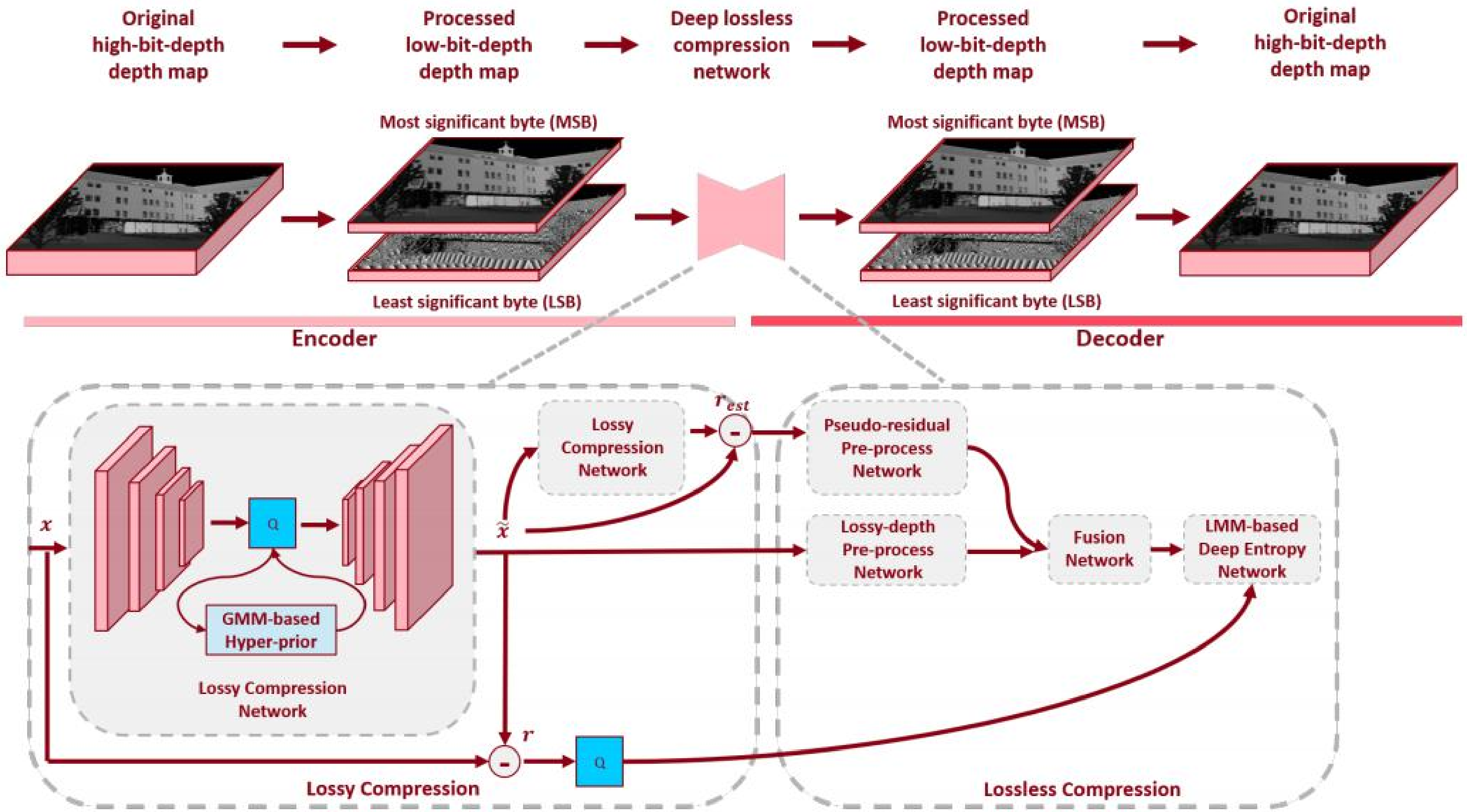


LiDAR : Velodyne HDL-64E LiDAR



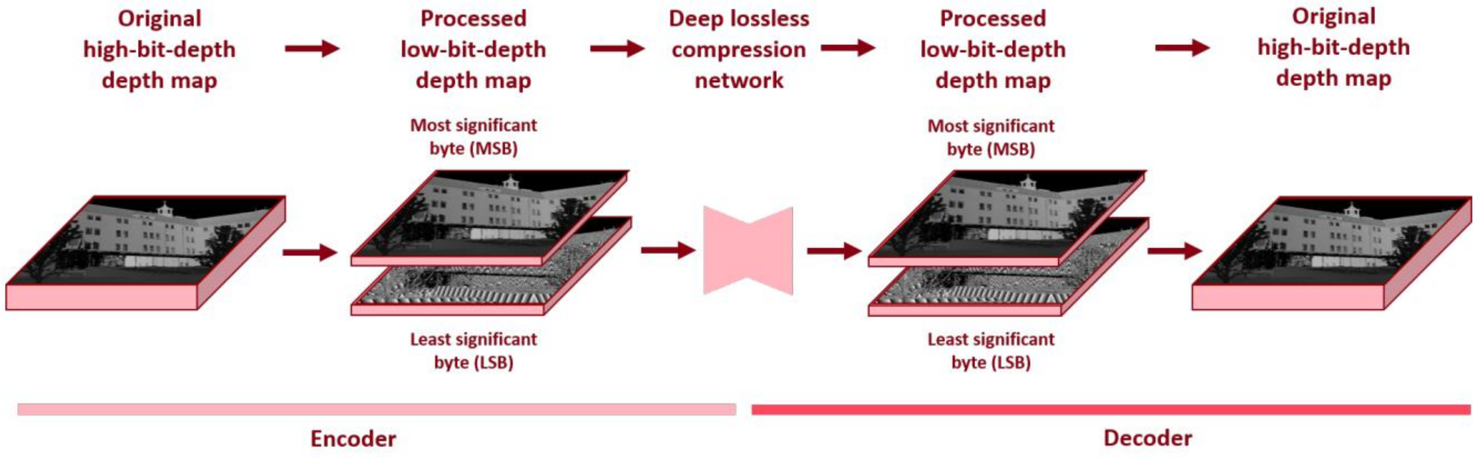
Proposed method

End-to-end lossless compression network targeted at the high precision depth maps



Pre-processing of depth maps

We split the high-bit-depth depth map into two low-bit-depth depth maps. These two low-bit-depth depth maps are denoted as **most significant bytes (MSB)** and **least significant bytes (LSB)** respectively.



Splitting

$$x_{MSB_i} = \left\lfloor \frac{x_i}{d} \right\rfloor$$

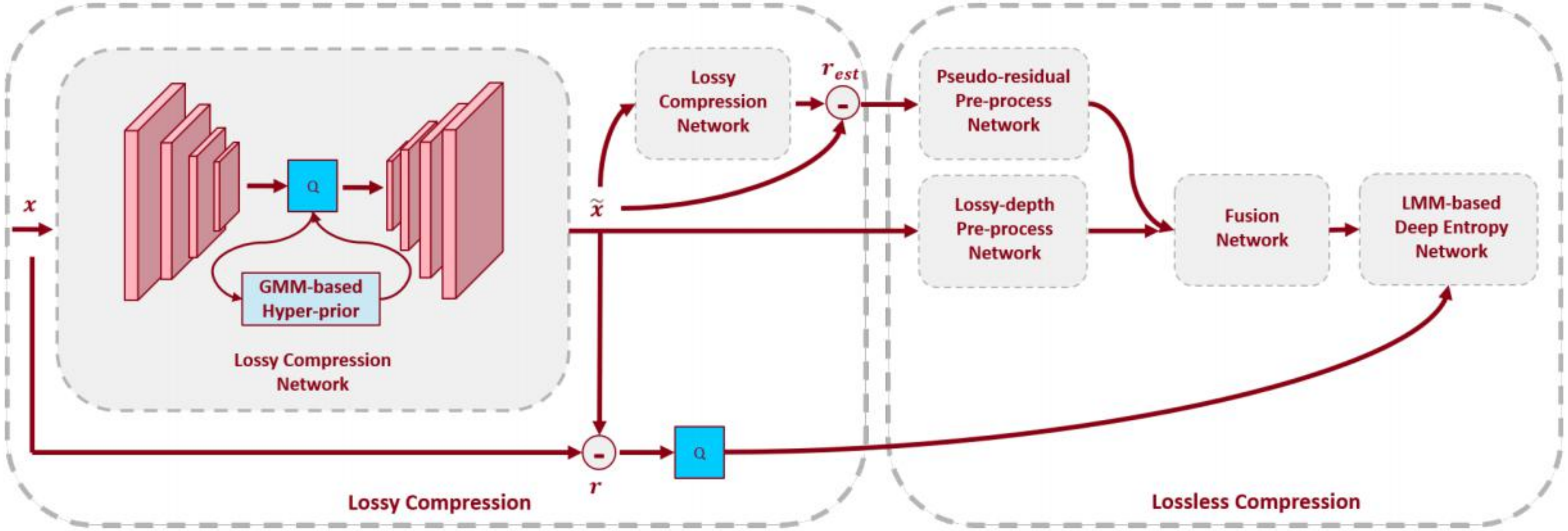
$$x_{LSB_i} = x_i \bmod d$$

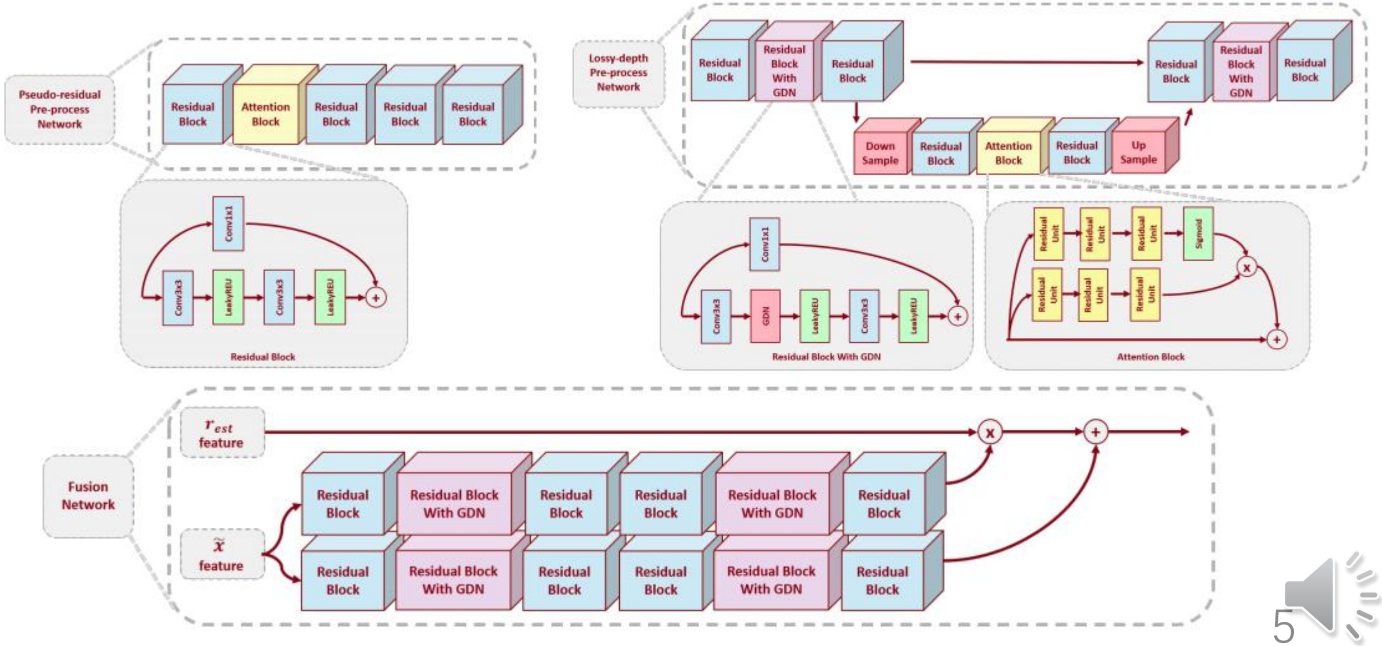
Recovering

$$x_i = x_{MSB_i} * d + x_{LSB_i}$$

Joint lossy depth maps & lossless residual compression

Pseudo-residual is used to guide the generation of distribution for real-residual.





$$\tilde{x} = C(x)$$

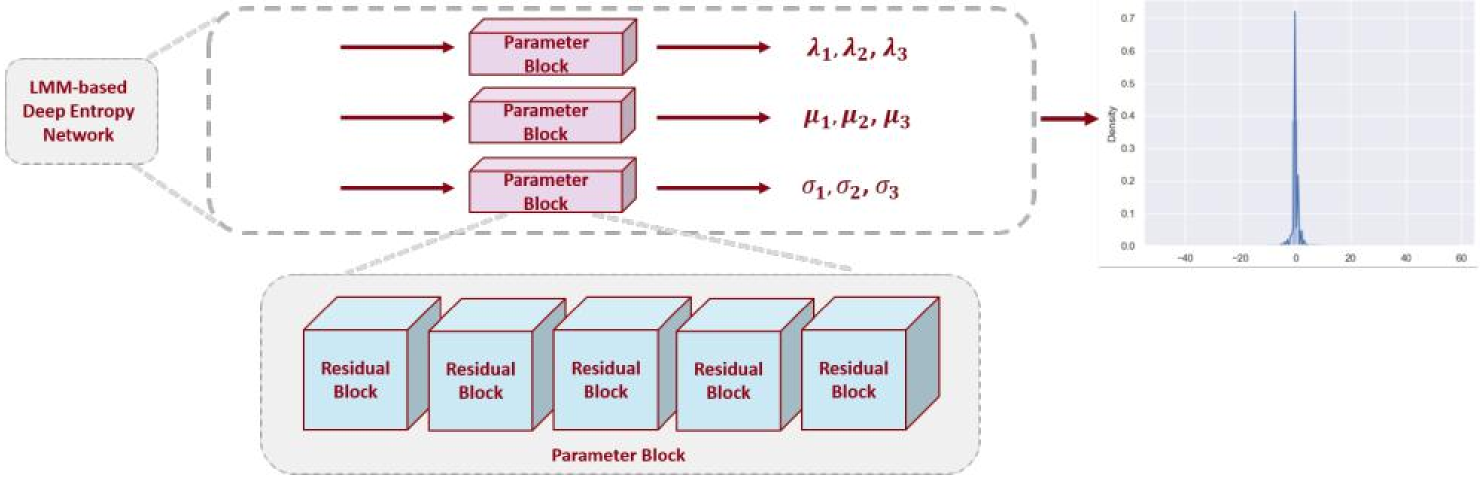
$$\tilde{x}_{sim} = C(\tilde{x})$$

Real-residual → $r = \mathrm{round}(x - \tilde{x})$

Pseudo-residual → $r_{est} = \tilde{x} - \tilde{x}_{sim}$

Joint lossy depth maps & lossless residual compression

We model the distribution of real-residual with a Laplace mixture distribution.



Single Laplace Distribution →

$$f(r|\sigma,\mu) = \frac{1}{2\sigma} e^{-\frac{|r-\mu|}{\sigma}}$$

Laplace Mixture Distribution →

$$P(r|\lambda,\sigma,\mu) = \sum_{k=1}^{K} \lambda_k f(r|\sigma_k,\mu_k), with \sum_{k=1}^{K} \lambda_k = 1$$

Loss function

Overall, there are two compression processes: the first one is a lossy compression and the other is a lossless compression. Therefore, the whole compression rate consists of two parts.

$$R_{lossy} = \mathbb{E}_{\tilde{p}(x)}\,\mathbb{E}_{q(\hat{y},\hat{z}|x)}[-\log_2 p(\hat{z}) - \log_2 p(\hat{y}|\hat{z})]$$

$$R_{lossless} = \mathbb{E}_{\tilde{p}(x)}\,\mathbb{E}_{q(\tilde{x},r_{est}|x)}[-\log_2 p(r|\tilde{x}, r_{est})]$$

Besides rate terms, we add two distortion terms.

$$D(x,\tilde{x}) = \mathbb{E}_{\tilde{p}(x)}\,\mathbb{E}_{q(\tilde{x}|x)}\,||x-\tilde{x}||_2$$

$$D(r,r_{est}) = \mathbb{E}_{\tilde{p}(x)}\,\mathbb{E}_{q(r,r_{est}|x)}\,||r-r_{est}||_2$$

$$L = R_{lossy} + R_{lossless} + \alpha * D(x,\tilde{x}) + \beta * D(r,r_{est})$$

Results

We compare our proposed method with nine non-learned codecs and two learned codecs, including lossless data compression methods and lossless image compression methods.

Table 1: Compression performance comparison of our method and eleven engineered codecs. We show lossless compression performance in terms of bpp (bits per pixel). The precision of depth maps is $1mm$.

coding schemes		DIODE	SementicKITTI
non-learned data lossless	ZLIB	8.4090	13.4663
	GZIP	7.9963	12.8679
	BZ2	4.9614	9.7833
	LZMA	4.9542	8.5063
non-learned image lossless	AVIF	6.3712	11.9195
	BPG	5.5042	10.0245
	PNG	5.2862	10.0981
	FILF	4.0304	10.8983
	WEBP	4.0807	7.9249
learned image lossless	IDF	5.4870	9.7110
	SReC	4.9668	8.9721
learned depth map lossless	Ours	**3.7514**	**7.3725**

Table 4: Ablation study on fusion network and LMM-based deep entropy network with $\alpha = 25$, $\beta = 25$, $d = 512$.

datasets	modules		R_{lossy}		$R_{lossless}$	overall bpp
	fusion network	LMM-based deep entropy network	$R_{\hat{y}}$	$R_{\hat{z}}$		
DIODE	✓	✓	0.3265	0.0103	3.4147	**3.7514**
		✓	0.3002	0.0101	3.5955	3.9059
	✓		0.3322	0.0103	3.5058	3.8483
			0.2516	0.0093	3.7428	4.0037

Sensitivity analysis

Sensitivity analysis for hyper-parameters.

Table 2: Sensitivity analysis for hyper parameter d with $\alpha = 100$ and $\beta = 100$.

datasets	hyper parameter setting	R_{lossy}		$R_{lossless}$	overall bpp
		$R_{\hat{y}}$	$R_{\hat{z}}$		
DIODE	$d = 1024$	0.3630	0.0100	3.5692	3.9423
	$d = 512$	0.3413	0.0101	3.4838	**3.8352**
	$d = 256$	0.3590	0.0103	3.4949	3.8642
	$d = 128$	0.3573	0.0103	3.5680	3.9357
	$d = 64$	0.3625	0.0099	3.6571	4.0295
SementicKITTI	$d = 128$	0.3734	0.0090	7.2628	7.6453
	$d = 64$	0.3683	0.0088	7.3465	7.7236
	$d = 32$	0.3813	0.0093	7.1916	7.5823
	$d = 16$	0.4245	0.0096	7.0674	**7.5015**
	$d = 8$	0.4913	0.0099	7.0735	7.5747

Table 3: Sensitivity analysis for hyper parameters α and β.

datasets	hyper parameters setting	R_{lossy}		$R_{lossless}$	overall bpp
		$R_{\hat{y}}$	$R_{\hat{z}}$		
DIODE	$\alpha = 400,\ \beta = 400,\ d = 512$	0.3681	0.0102	3.5198	3.8981
	$\alpha = 200,\ \beta = 200,\ d = 512$	0.3493	0.0102	3.4500	3.8096
	$\alpha = 100,\ \beta = 100,\ d = 512$	0.3413	0.0101	3.4838	3.8352
	$\alpha = 50,\ \beta = 50,\ d = 512$	0.3316	0.0103	3.4251	3.7669
	$\alpha = 25,\ \beta = 25,\ d = 512$	0.3265	0.0103	3.4147	**3.7514**
SementicKITTI	$\alpha = 400,\ \beta = 400,\ d = 16$	0.4840	0.0100	7.1261	7.6201
	$\alpha = 200,\ \beta = 200,\ d = 16$	0.4533	0.0099	7.0739	7.5370
	$\alpha = 100,\ \beta = 100,\ d = 16$	0.4245	0.0096	7.0674	7.5015
	$\alpha = 50,\ \beta = 50,\ d = 512\ d = 16$	0.4112	0.0097	7.0195	7.4405
	$\alpha = 25,\ \beta = 25,\ d = 512\ d = 16$	0.4004	0.0097	6.9624	**7.3725**

Thank you for watching!

End-to-end lossless compression of high precision depth maps guided by pseudo-residual
DCC 2022